

# 一种基于 PROSPECTOR 及 MYCIN 系统的综合推理模型 A Comprehensive Reasoning Model Based on PROSPECTOR System and MYCIN System

陈 军      陈 严\*  
Chen Jun    Chen Yan

(广西财政高等专科学校 南宁市明秀西路 100 号 530003)  
(Guangxi Training School of Public Finance,  
100 West Mingxiulu, Nanning, Guangxi, 530003)

**摘要** 分析 MYCIN 的可信度  $CF$  与 PROSPECTOR 的充分性度量  $LS$  的关系. 将可信度  $CF$  引入到主观 Bayes 模型上, 导出一个综合的推理模型. 该模型不仅消除了  $CF$  与  $LS$  间可能存在的 inconsistency, 而且减少了要求先验概率值的结点数.

**关键词** 不确定性推理 主观 Bayes 可信度 专家系统  
中图法分类号 TP 18

**Abstract** The relation between MYCIN certainty factor ( $CF$ ) and PROSPECTOR sufficiency measure ( $LS$ ) is analysed. By introducing  $CF$  into subjective Bayes reasoning model, an improved model is derived. This model not only eliminates the possible inconsistency between  $CF$  and  $LS$ , but also reduces the necessary nodes of prior probability.

**Key words** uncertainty reasoning, subjective Bayes, certainty factor, expert system

不确定性信息的处理是专家系统要解决的主要问题之一. 事实上, 不确定性是人们思维中经常出现的一种心理状态, 现实世界也没有完全确定的事情. 如何表示和处理这些不确定性就成为人工智能研究的一项重要内容, 研究不确定性推理技术是提高专家系统性能的有效途径之一. 1976 年 Duda 等根据 PROSPECTOR 的推理模型建立了主观 Bayes 的不确定推理模型, 引入两个值 ( $LS$  和  $LN$ ) 来度量不确定性. 其中  $LS$  表现规则成立的充分性,  $LN$  表现规则成立的必要性. 这种表示的优点是既考虑了证据  $E$  的出现对结论  $H$  的支持, 又考虑了证据  $E$  的不出现对结论  $H$  的影响, 在推理计算方面具有坚实的概率理论为基础. 但是, 这种方法的缺点是在实际应用中要求人类专家给推理网络中的所有结点赋多个先验概率值, 这是极难做到的, 同时也不符合人类推理的习惯. 因此, 很少有专家系统使用这一方法. 另一著名的 MYCIN 系统, 采用确定因子或称可信度 ( $CF$ ) 度量不确定性.  $CF$  既表现对证据的信

任程度, 又体现因证据  $E$  出现而引起对结论  $H$  的信任程度, 因此在实际使用中人类专家不必给推理网络上的结点赋先验概率. 但这一系统的缺点是在推理计算方面缺乏良好的理论基础. 另一方面,  $LS$  和  $CF$  由人类专家给出, 这可能导致它们之间的 inconsistency.

本文综合 PROSPECTOR 与 MYCIN 二者的优点而导出一种综合的推理模型, 从而建立了  $CF$  与  $LS$ 、 $LN$  的关系, 并能消除它们之间的 inconsistency.

## 1 可信度 $CF$ 与充分性度量 $LS$ 的关系

尽管在 MYCIN 和 PROSPECTOR 系统中  $CF$ 、 $LS$  和  $LN$  均由专家主观给出, 但它们都给出了基于概率的解释.

### 1.1 $CF$ 的意义

在 MYCIN 系统中, 对  $CF$  做了如下的规定<sup>[3]</sup>:

对于规则  $R: E \rightarrow H$  有:

$$CF(H, E) = \begin{cases} \frac{P(H|E) - P(H)}{1 - P(H)} & \text{当 } P(H|E) \geq P(H) \\ \frac{P(H|E) - P(H)}{P(H)} & \text{当 } P(H|E) < P(H) \end{cases} \quad (1)$$

其中  $P$  表示概率,  $CF(H, E)$  的意义为:

1)  $CF(H, E) = 1$  时,  $P(H|E) = 1$  表示证据  $E$

1998-02-09 收稿, 1998-03-13 修回.

\* 广西区党委宣传部, 南宁市民族大道, 530023 (Guangxi Propaganda Department, Minzhu Dadao, Nanning, Guangxi, 530023).

真, 结论  $H$  必真;

2)  $CF(H, E) = 0$  时,  $P(H | E) = P(H)$  表示证据  $E$  与结论  $H$  无关;

3)  $CF(H, E) = -1$  时,  $P(H | E) = 0$  表示证据  $E$  真, 结论  $H$  必假;

4)  $CF(H, E) > 0$  时, 表示证据  $E$  真支持结论  $H$  真;

5)  $CF(H, E) < 0$  时, 表示证据  $E$  真不支持结论  $H$  真.

### 1.2 证据的不确定性度量

1)  $CF(E) = 1$  时, 表示证据  $E$  为真;

2)  $CF(E) = 0$  时, 表示对证据  $E$  一无所知;

3)  $CF(E) = -1$  时, 表示证据  $E$  为假;

4)  $CF(E) > 0$  时, 表示证据  $E$  以  $CF(E)$  程度为真;

5)  $CF(E) < 0$  时, 表示证据  $E$  以  $CF(E)$  程度为假.

### 1.3 $LS$ 和 $LN$ 的概率解释

在 PROSPECTOR 系统中, 文献[3] 给出了  $LS$  和  $LN$  的概率解释和定义:

$$LS = \frac{P(E | H)}{P(E | \sim H)}$$

$$LN = \frac{P(\sim E | H)}{P(\sim E | \sim H)} = \frac{1 - P(E | H)}{1 - P(E | \sim H)}$$

并给出主观 Bayes 推理模型:

$$\begin{cases} P(H | E) = \frac{LS \cdot P(H)}{(LS - 1) \cdot P(H) + 1} \\ P(H | \sim E) = \frac{LN \cdot P(H)}{(LN - 1) \cdot P(H) + 1} \end{cases} \quad (2)$$

从(2)可知,  $LS$  与  $LN$  是对等的, 且分别表示  $E$  和  $\sim E$  的出现对  $H$  的支持程度.

$LS$  的意义:

1)  $LS \rightarrow +\infty$  时, 说明  $E$  的存在导致  $H$  为真;

2)  $LS = 1$  时, 说明  $E$  与  $H$  无关;

3)  $LS > 1$  时, 说明  $E$  的存在支持  $H$  的存在;

4)  $LS < 1$  时, 说明  $E$  的存在使  $H$  存在的可能性下降;

5)  $LS = 0$  时, 说明  $E$  的存在使  $H$  为假.

### 1.4 建立 $CF$ 与 $LS$ 的联系

由 (2) 得:

$$LS = \frac{P(H | E) \cdot (P(H) - 1)}{P(H) \cdot (P(H | E) - 1)}$$

$$\frac{1}{LS} = \frac{P(H) \cdot (P(H | E) - 1)}{(P(H) - 1) \cdot P(H | E)} = \frac{-P(H)}{P(H | E)}$$

$$\left( \frac{P(H | E) - P(H)}{1 - P(H)} - 1 \right)$$

当  $P(H | E) \geq P(H)$  时, 由(1)得:

$$\frac{1}{LS} = \frac{-P(H)}{P(H | E)} \cdot (CF(H, E) - 1)$$

$$LS = \frac{P(H | E)}{P(H)} \cdot \frac{1}{1 - CF(H, E)}$$

同理, 当  $P(H | E) < P(H)$  时得:

$$LS = \frac{1 - P(H)}{1 - P(H | E)} \cdot (CF(H, E) + 1)$$

即:

$$LS = \begin{cases} \frac{P(H | E)}{P(H)} \cdot \frac{1}{1 - CF(H, E)} & \text{当 } 0 \leq CF(H, E) \leq 1 \\ \frac{1 - P(H)}{1 - P(H | E)} \cdot (CF(H, E) + 1) & \text{当 } -1 \leq CF(H, E) < 0 \end{cases} \quad (3)$$

(3) 式反映了  $CF$  与  $LS$  的关系, 并且:

1)  $CF(H, E) = 0$ ,  $E$  与  $H$  无关, 得  $LS = 1$ ;

2)  $CF(H, E) = 1$ ,  $E$  真使  $H$  真, 得  $LS \rightarrow +\infty$ ;

3)  $CF(H, E) = -1$ ,  $E$  真使  $H$  假, 得  $LS = 0$ ;

4)  $CF(H, E) > 0$ ,  $E$  支持  $H$ , 得  $LS > 1$ ;

5)  $CF(H, E) < 0$ ,  $E$  不支持  $H$ , 得  $LS < 1$ .

可见, 可信度  $CF$  与充分性度量  $LS$  所表示的涵义是相容的.

## 2 建立推理模型

前面已经建立了可信度  $CF$  与充分性度量  $LS$  存在着式(3)的关系, 并且它们所表示的涵义是相容的. 这样就有理由把可信度  $CF$  因子引入到主观 Bayes 模型上去.

当  $P(H | E) \geq P(H)$  时, 由(3)得:

$$P(H) = P(H | E) / (LS \cdot (1 - CF(H, E)))$$

并代入(2)得:

$$\begin{aligned} P(H | E) &= \frac{LS \cdot P(H)}{1 + (LS - 1) \cdot P(H)} \\ &= \frac{LS \cdot P(H | E) / (LS \cdot (1 - CF(H, E)))}{1 + (LS - 1) \cdot P(H | E) / (LS \cdot (1 - CF(H, E)))} \\ &= \frac{LS \cdot P(H | E)}{(LS - 1) \cdot P(H | E) + LS \cdot (1 - CF(H, E))} \end{aligned}$$

经整理得:

$$P(H | E) \cdot (LS - 1) = LS \cdot CF(H, E)$$

$$P(H | E) = \frac{LS \cdot CF(H, E)}{LS - 1}$$

同理可得, 当  $P(H | E) < P(H)$  时:

$$P(H | E) = 1 + \frac{CF(H, E)}{1 - LS}$$

故得推理模型:

$$P(H | E) = \begin{cases} \frac{LS \cdot CF(H, E)}{LS - 1} & \text{当 } 0 \leq CF(H, E) \leq 1 \\ 1 + \frac{CF(H, E)}{1 - LS} & \text{当 } -1 \leq CF(H, E) < 0 \end{cases} \quad (4)$$

由于  $LS$  与  $LN$  有对等的地位, 类似地有:

$$P(H | \sim E) = \begin{cases} \frac{LN \cdot CF(H, \sim E)}{LN - 1} & \text{当 } 0 \leq CF(H, \sim E) \leq 1 \\ 1 + \frac{CF(H, \sim E)}{1 - LN} > & \text{当 } -1 \leq CF(H, \sim E) < 0 \end{cases} \quad (5)$$

从以上两式可以看出, 在求  $P(H | E)$  或  $P(H | \sim E)$  时, 并没有涉及到先验概率值(如  $P(H)$ ), 这对人类专家来说给出  $CF$  和  $LS$ 、 $LN$  值要比给出先验概率要容易得多。由于  $CF$ 、 $LS$  和  $LN$  是表示不确定性的度量, 故更符合人类的思维习惯。

下面讨论结论的可信度计算

(i) 证据是确定时

1) 当  $P(E) = 1$  时, 直接应用式(4) 求出  $P(H | E)$  值, 可信度为:

$$CF(H) = CF(H, E);$$

2) 当  $P(E) = 0$  时, 直接应用式(5) 求出  $P(H | \sim E)$  值, 可信度为:

$$CF(H) = CF(H, \sim E).$$

(ii) 证据为不确定的, 即  $0 < P(E) < 1$  时

设  $E'$  代表与  $E$  有关的所有观察, 则有:

$$P(H | E') = P(H | E) \cdot P(E | E') + P(H | \sim E) \cdot P(\sim E | E')$$

1) 当  $P(E | E') = 1$  时, 证据必然出现

$$P(H | E') = P(H | E) =$$

$$\begin{cases} \frac{LS \cdot CF(H, E)}{LS - 1} & \text{当 } 0 \leq CF(H, E) \leq 1 \\ 1 + \frac{CF(H, E)}{1 - LS} & \text{当 } -1 \leq CF(H, E) < 0 \end{cases}$$

$$CF(H) = CF(H, E);$$

2) 当  $P(E | E') = 0$  时,

$$P(H | E') = P(H | \sim E) =$$

$$\begin{cases} \frac{LN \cdot CF(H, \sim E)}{LN - 1} & \text{当 } 0 \leq CF(H, \sim E) \leq 1 \\ 1 + \frac{CF(H, \sim E)}{1 - LN} & \text{当 } -1 \leq CF(H, \sim E) < 0 \end{cases}$$

$$CF(H) = CF(H, \sim E)$$

3) 当  $P(E | E') = P(E)$  时,  $P(H | E') = P(H)$

4) 当  $P(E | E')$  为其它值时, 可以采用 MYCIN 的推理方法:

(a) 证据为单个条件时

$$P(H | E) =$$

$$\begin{cases} \frac{LS \cdot CF(H, E)}{LS - 1} & \text{当 } 0 \leq CF(H, E) \leq 1 \\ 1 + \frac{CF(H, E)}{1 - LS} & \text{当 } -1 \leq CF(H, E) < 0 \end{cases}$$

$$CF(H) = CF(H, E) \cdot \max\{0, CF(E)\}$$

其中  $CF(E)$  由专家主观给出。

(b) 证据是多个条件的逻辑组合

若  $E_1 \wedge E_2 \wedge \dots \wedge E_n \rightarrow H(x)$ , 则

$$CF(E) = CF(E_1 \wedge E_2 \wedge \dots \wedge E_n) = \min\{CF(E_1), CF(E_2), \dots, CF(E_n)\}$$

若  $E = E_1 \vee E_2 \vee \dots \vee E_n$ , 则

$$CF(E) = CF(E_1 \vee E_2 \vee \dots \vee E_n) = \max\{CF(E_1), CF(E_2), \dots, CF(E_n)\}$$

当两条规则具有相同结论

$$E_1 \rightarrow H, \quad CF(H, E_1)$$

$$E_2 \rightarrow H, \quad CF(H, E_2),$$

$$\text{则 } CF_1(H) = CF(H, E_1) \cdot \max\{0, CF(E_1)\}$$

$$CF_2(H) = CF(H, E_2) \cdot \max\{0, CF(E_2)\}$$

然后用公式:

$$CF_{12}(H) =$$

$$\begin{cases} CF_1(H) + CF_2(H) - CF_1(H) \cdot CF_2(H) & \text{当 } CF_1(H) \& CF_2(H) \geq 0 \\ CF_1(H) + CF_2(H) + CF_1(H) \cdot CF_2(H) & \text{当 } CF_1(H) \& CF_2(H) < 0 \\ CF_1(H) + CF_2(H) & \text{其它} \end{cases}$$

### 3 讨论

本文从分析可信度  $CF$  与充分性度量  $LS$  的关系着手, 把 MYCIN 的可信度  $CF$  引入到主观 Bayes 模型中, 从而得出了综合的推理模型:

$$P(H | E) =$$

$$\begin{cases} \frac{LS \cdot CF(H, E)}{LS - 1} & \text{当 } 0 \leq CF(H, E) \leq 1 \\ 1 + \frac{CF(H, E)}{1 - LS} & \text{当 } -1 \leq CF(H, E) < 0 \end{cases}$$

$$P(H | \sim E) =$$

$$\begin{cases} \frac{LN \cdot CF(H, \sim E)}{LN - 1} & \text{当 } 0 \leq CF(H, \sim E) \leq 1 \\ 1 + \frac{CF(H, \sim E)}{1 - LN} & \text{当 } -1 \leq CF(H, \sim E) < 0 \end{cases}$$

由于用  $CF$  和  $LS$  来表示先验概率  $P(H)$ , 使得这个推理模型不需要先验概率。在推理计算上对于具有确定性的证据, 仍采用主观 Bayes 的方法; 在对于具有不确定性的证据时, 采用主观 Bayes 方法和 MYCIN 方法相结合的办法, 使得计算结果为两个数值, 一个是反映结论发生的概率  $P(H | E)$ ; 另一个是反映结论的可信任程度  $CF(H)$ 。换句话说, 只有当结果具有较高的发生概率和较高的可信度时, 才是正确的。本模型既遗传了 PROSPECTOR 具有良好的理论基础, 又采纳了 MYCIN 不需要先验概率的优点, 并能兼容 MYCIN 的确定因子型规则强度和 PROSPECTOR 的充分性度量和必要性度量规则强度, 以及它们的组合形式。

(下转第 200 页 Continue on page 200)

对象在缓冲区的频繁换入换出, 不可避免地会出现一些存储碎片和无用的对象。为此, 系统存储管理器设置了一个回收模块, 定期回收一些暂时不用的缓冲区空间和 COAT 表项。回收算法采用 LRU 算法。

### 3.4 用户接口

用户接口是为工程设计应用软件提供的一套接口函数, 供应用程序建立、查询、修改数据库时调用。在 REDBMS 系统中, 配置了一套简洁方便、功能比较齐全的用户接口, 其功能有: (1) 设计的打开、初始化、删除、关闭; (2) 设计版本的创建、打开、删除; (3) 表结构的建立、初始化、删除及查询下一级子表; (4) 以常数、内存变量为操作对象的 I/O 操作。包括增加一条记录中的数据、替换一条记录中的数据、替换全部记录中的数据、条件替换一条记录中的数据、条件替换所有记录中的数据、读一条记录的数据等; (5) 以结构为操作对象的 I/O 操作。增加一个结构数组、替换一个结构数组、条件替换一个结构数组、读数据到一个结构数组中等; (6) 以数组为操作对象的 I/O 操作。增加一个数组、替换一个数组、条件替换一个数组、读数据到一个数组中等; (7) 数据记录的插入、查询、删除、恢复和压缩; (8) 统计、计算; (9) 库文件结构的查询; (10) 索引文件的建立、查询和释放。

## 4 结语

我们运用集成技术和面向对象方法, 在一个统一的集成 CAD 环境下开发了一个集成式工程数据库管理系统 REDBMS。该系统有效地解决了铁路中间

站站场设计过程中工程数据的描述、操作和存储管理等问题, 有力地支持了 WSCADP 系统的集成。WSCADP 系统投入实际应用后, 实践也表明 REDBMS 系统的集成方法和存储管理是非常有效的, 满足了站场设计对工程数据存储管理的需要, 其性能完全符合设计要求。

### 参考文献

- 1 李陶深, 张龙祥. 工程数据库中复杂对象的存储管理. 长沙铁道学院学报, 1991, 9 (1): 21~28.
- 2 寿宇澄, 何志均, 董金祥. OSCAR II: 面向对象的工程数据库管理系统. 计算机辅助设计与图形学学报, 1994, 6 (2): 143~150.
- 3 俞勇, 魏国臣, 黄上腾. 基于面向对象工程数据库的 CAD/CAPP 系统的集成. 计算机工程, 1996, 22 (3): 32~36.
- 4 吴弘静, 林中尧, 刘大昕. 面向对象的工程数据库模型. 小型微型计算机系统, 1996, 17 (3): 22~25.
- 5 张龙祥, 李陶深, 张学丽. 微机 CAD 系统实用开发技术. 武汉: 湖北科学技术出版社, 1994. 12.
- 6 顾宁, 林宗楷, 郭玉钊. EDBMS/3: 一个微机环境下的工程数据库管理系统. 计算机辅助设计与图形学学报, 1995, 7 (3): 213~219.
- 7 蔡青, 高光焘. CAD/CAM 系统的可视化、集成化、智能化、网络化. 西安: 西北工业大学出版社, 1996. 11.
- 8 李韶原, 罗晓沛. 数据库技术新进展. 北京: 清华大学出版社, 1997. 5.
- 9 Dittrich W. On object-oriented database systems, Springer-Verlay, 1991.
- 10 Selinger P G. Prediction and challenges for database systems in the years 2000. In: Proc of the 19th VLDB Conference, 1993. 667~675. (责任编辑: 黎贞崇)

(上接第 196 页 Continue from page 196)

### 致谢

本文得到了广西师范大学张师超教授的热心关心和指导, 在此表示衷心的感谢!

### 参考文献

- 1 何新贵. 知识处理与专家系统. 北京: 国防工业出版社, 1990. 9.
- 2 张师超, 严小卫, 王成名等. 不确定性推理技术. 桂林: 广西师范大学出版社, 1996.

- 3 石纯一, 黄昌宁, 王家等. 人工智能原理. 北京: 清华大学出版社, 1993.
- 4 Chen S et al. . An inexact reasoning algorithm based on fuzzy rule matrix transformation. J of Automated Reasoning, 1992. 8.
- 5 Smith E, Langston C, Niobett R. The case for rules in reasoning. Cognitive Science, 1992. 16: 295~355.
- 6 Zadeh L. Calculus of fuzzy restriction, fuzzy: Sets and Their Application to Cognitive Decision Processes New York: Academic Press, 1995.
- 7 Barletta R. An introduction to case-based reasoning, AI-expert, 1991, 8: 43~49.

(责任编辑: 黎贞崇)